

Multiplayer Cost Games with Simple Nash Equilibria

Thomas Brihaye¹, Julie De Pril¹, and Sven Schewe²

¹ University of Mons - UMONS
Place du Parc 20, 7000 Mons, Belgium
{thomas.brihaye, julie.depril}@umons.ac.be

² University of Liverpool
sven.schewe@liverpool.ac.uk

Abstract. Multiplayer games with selfish agents naturally occur in the design of distributed and embedded systems. As the goals of selfish agents are usually neither equivalent nor antagonistic to each other, such games are non zero-sum games. We study such games and show that a large class of these games, including games where the individual objectives are mean- or discounted-payoff, or quantitative reachability, and show that they do not only have a solution, but a *simple* solution. We establish the existence of Nash equilibria that are composed of k memoryless strategies for each agent in a setting with k agents, one main and $k - 1$ minor strategies. The main strategy describes what happens when all agents comply, whereas the minor strategies ensure that all other agents immediately start to co-operate against the agent who first deviates from the plan. This simplicity is important, as rational agents are an idealisation. Realistically, agents have to decide on their moves with very limited resources, and complicated strategies that require exponential—or even non-elementary—implementations cannot realistically be implemented. The existence of simple strategies that we prove in this paper therefore holds a promise of implementability.

1 Introduction

The construction of correct and efficient computer systems (both hard- and software) is recognised to be an extremely difficult task. Formal methods have been exploited with some success in the design and verification of such systems. Mathematical logic, automata theory [17], and model-checking [12] have contributed much to the success of formal methods in this field. However, traditional approaches aim at systems with qualitative specifications like LTL, and rely on the fact that these specifications are either satisfied or violated by the system.

Unfortunately, these techniques do not trivially extend to complex systems, such as embedded or distributed systems. A main reason for this is that such systems often consist of multiple independent components with individual objectives. These components can be viewed as selfish agents that may cooperate and compete at the same time. It is difficult to model the interplay between these components with traditional finite state machines, as they cannot reflect the intricate quantitative valuation of an agent on how well he has met his goal. In particular, it is not realistic to assume that these components

are always cooperating to satisfy a common goal, as it is, e.g., assumed in works that distinguish between an environment and a system. We argue that it is more realistic to assume that all components act like selfish agents that try to achieve their own objectives and are either unconcerned about the effect this has on the other components or consider this effect to be secondary. It is indeed a recent trend to enhance the system models used in the classical approach of verification by quantitative cost and gain functions, and to exploit the well established game-theoretic framework [21,22] for their formal analysis.

The first steps towards the extension of computational models with concepts from classical game theory were taken by advancing from boolean to general two-player zero-sum games played on graphs [15]. Like their qualitative counter parts, those games are adequate to model controller-environment interaction problems [24,25]. As usual in control theory, one can distinguish between moves of a control player, who plays actions to control a system to meet a control objective, and an antagonistic environment player. In the classical setting, the control player has a qualitative objective—he might, for example, try to enforce a temporal specification—whereas the environment tries to prevent this. In the extension to quantitative games, the controller instead tries to maximise its gain, while the environment tries to minimise it. This extension lifts the controller synthesis problem from a constructive extension of a decision problem to a classical optimisation problem.

However, this extension has not lifted the restriction to purely antagonist interactions between a controller and a hostile environment. In order to study more complex systems with more than two components, and with objectives that are not necessarily antagonist, we resort to multiplayer non zero-sum games. In this context, *Nash equilibria* [21] take the place that winning and optimal strategies take in qualitative and quantitative two-player games zero-sum games, respectively. Surprisingly, qualitative objectives have so far prevailed in the study of Nash equilibria for distributed systems. However, we argue that Nash equilibria for selfish agents with quantitative objectives—such as reaching a set of target states quickly or with a minimal consumption of energy—are natural objectives that ought to be studied alongside (or instead of) traditional qualitative objectives.

Consequently, we study *Nash equilibria* for *multiplayer non zero-sum* games played on graphs with *quantitative* objectives.

Our contribution. In this paper, we study turn-based multiplayer non zero-sum games played on finite graphs with quantitative objectives, expressed through a cost function for each player (*cost games*). Each cost function assigns, for every play of the game, a value that represents the cost that is incurred for a player by this play. Cost functions allow to express classical quantitative objectives such as *quantitative reachability* (i.e., the player aims at reaching a subset of states as soon as possible), or *mean-payoff* objectives. In this framework, all players are supposed to be rational: they want to minimise their own cost or, equivalently, maximise their own gain. This invites the use of Nash equilibria as the adequate concept for cost games.

Our results are twofold. Firstly, we prove the *existence* of Nash equilibria for a large class of cost games that includes quantitative reachability and mean-payoff objectives. Secondly, we study the complexity of these Nash equilibria in terms of the *memory*

needed in the strategies of the individual players in these Nash equilibria. More precisely, we ensure existence of Nash equilibria whose strategies only requires a number of memory states that is *linear* in the size of the game for a wide class of cost games, including games with quantitative reachability and mean-payoff objectives.

The general philosophy of our work is as follows: we try to derive existence of Nash equilibria in multiplayer non zero-sum quantitative games (and characterization of their complexity) through determinacy results (and characterization of the optimal strategies) of several well-chosen two-player quantitative games derived from the multiplayer game. These ideas were already successfully exploited in the qualitative framework [16], and in the case of limit-average objectives [26].

Related work. Several recent papers have considered *two-player zero-sum* games played on finite graphs with regular objectives enriched by some *quantitative* aspects. Let us mention some of them: games with finitary objectives [10], mean-payoff parity games [11], games with prioritised requirements [1], request-response games where the waiting times between the requests and the responses are minimized [18,28], games whose winning conditions are expressed via quantitative languages [2], and recently, cost-parity and cost-Streett games [13].

Other work concerns *qualitative non zero-sum* games. In [16], general criteria ensuring existence of Nash equilibria and subgame perfect equilibria (resp. secure equilibria) are provided for multiplayer (resp. 2-player) games, as well as complexity results. The complexity of Nash equilibria in multiplayer concurrent games with Büchi objectives has been discussed in [5]. [4] studies the existence of Nash equilibria for timed games with qualitative reachability objectives

Finally, there is a series of recent results on the combination of *non zero-sum* aspects with *quantitative objectives*. In [3], the authors study games played on graphs with terminal vertices where quantitative payoffs are assigned to the players. In [19], the authors provide an algorithm to decide the existence of Nash equilibria for concurrent priced games with quantitative reachability objectives. In [23], the authors prove existence of a Nash equilibrium in Muller games on finite graphs where players have a preference ordering on the sets of the Muller table. Let us also notice that the existence of a Nash equilibrium in cost games with quantitative reachability objectives we study in this paper has already been established in [7]. The new proves we provide are simpler and significantly improve the complexity of the strategies constructed from exponential to linear in the size of the game.

Organization of the paper. In Section 2, we present the model of multiplayer cost games and define the problems we study. The main results are given in Section 3. Finally, in Section 4, we apply our general result on particular cost games with classical objectives. Omitted proofs and additional materials can be found in the Appendix.

2 General Background

In this section, we define our model of *multiplayer cost game*, recall the concept of Nash equilibrium and state the problems we study.

Definition 1. A multiplayer cost game is a tuple $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, E, (\text{Cost}_i)_{i \in \Pi})$ where

- Π is a finite set of players,
- $G = (V, E)$ is a finite directed graph with vertices V and edges $E \subseteq V \times V$,
- $(V_i)_{i \in \Pi}$ is a partition of V such that V_i is the set of vertices controlled by player i , and
- $\text{Cost}_i : \text{Plays} \rightarrow \mathbb{R} \cup \{+\infty, -\infty\}$ is the cost function of player i , where Plays is the set of plays in \mathcal{G} , i.e. the set of infinite paths through G . For every play $\rho \in \text{Plays}$, the value $\text{Cost}_i(\rho)$ represents the amount that player i loses for this play.

Cost games are *multiplayer turn-based quantitative non zero-sum* games. We assume that the players are rational: they play in a way to minimise their own cost.

Note that minimising cost or maximising gain are essentially³ equivalent, as maximising the gain for player i can be modelled by using Cost_i to be minus this gain and then minimising the cost. This is particularly important in cases where two players have antagonistic goals, as it is the case in all two-player zero-sum games. To cover these cases without changing the setting, we sometimes refer to maximisation in order to preserve the connection to such games in the literature.

For the sake of simplicity, we assume that each vertex has at least one outgoing edge. Moreover, it is sometimes convenient to specify an initial vertex $v_0 \in V$ of the game. We then call the pair (\mathcal{G}, v_0) an *initialised multiplayer cost game*. This game is played as follows. First, a token is placed on the initial vertex v_0 . Whenever a token is on a vertex $v \in V_i$ controlled by player i , player i chooses one of the outgoing edges $(v, v') \in E$ and moves the token along this edge to v' . This way, the players together determine an *infinite* path through the graph G , which we call a *play*. Let us remind that Plays is the set of all plays in \mathcal{G} .

A *history* h of \mathcal{G} is a *finite* path through the graph G . We denote by Hist the set of histories of a game, and by ϵ the empty history. In the sequel, we write $h = h_0 \dots h_k$, where $h_0, \dots, h_k \in V$ ($k \in \mathbb{N}$), for a history h , and similarly, $\rho = \rho_0 \rho_1 \dots$, where $\rho_0, \rho_1, \dots \in V$, for a play ρ . A *prefix* of length $n + 1$ (for some $n \in \mathbb{N}$) of a play $\rho = \rho_0 \rho_1 \dots$ is the finite history $\rho_0 \dots \rho_n$. We denote this history by $\rho[0, n]$.

Given a history $h = h_0 \dots h_k$ and a vertex v such that $(h_k, v) \in E$, we denote by hv the history $h_0 \dots h_k v$. Moreover, given a history $h = h_0 \dots h_k$ and a play $\rho = \rho_0 \rho_1 \dots$ such that $(h_k, \rho_0) \in E$, we denote by $h\rho$ the play $h_0 \dots h_k \rho_0 \rho_1 \dots$.

The function Last (resp. First) returns, for a given history $h = h_0 \dots h_k$, the last vertex h_k (resp. the first vertex h_0) of h . The function First naturally extends to plays.

A *strategy* of player i in \mathcal{G} is a function $\sigma : \text{Hist} \rightarrow V$ assigning to each history $h \in \text{Hist}$ that ends in a vertex $\text{Last}(h) \in V_i$ controlled by player i , a successor $v = \sigma(h)$ of $\text{Last}(h)$. That is, $(\text{Last}(h), \sigma(h)) \in E$. We say that a play $\rho = \rho_0 \rho_1 \dots$ of \mathcal{G} is *consistent* with a strategy σ of player i if $\rho_{k+1} = \sigma(\rho_0 \dots \rho_k)$ for all $k \in \mathbb{N}$ such that $\rho_k \in V_i$. A *strategy profile* of \mathcal{G} is a tuple $(\sigma_i)_{i \in \Pi}$ of strategies, where σ_i refers to a strategy for player i . Given an initial vertex v , a strategy profile determines the unique play of (\mathcal{G}, v) that is consistent with all strategies σ_i . This play is called the *outcome* of $(\sigma_i)_{i \in \Pi}$ and denoted by $\langle (\sigma_i)_{i \in \Pi} \rangle_v$. We say that a player *deviates* from a strategy (resp. from a play) if he does not carefully follow this strategy (resp. this play).

³ Sometimes the translation implies minor follow-up changes, e.g., the replacement of \liminf by \limsup and vice versa.

A *finite strategy automaton* for player $i \in \Pi$ over a game $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, E, (\text{Cost}_i)_{i \in \Pi})$ is a Mealy automaton $\mathcal{A}_i = (M, m_0, V, \delta, \nu)$ where:

- M is a non-empty, finite set of memory states,
- $m_0 \in M$ is the initial memory state,
- $\delta : M \times V \rightarrow M$ is the memory update function,
- $\nu : M \times V_i \rightarrow V$ is the transition choice function, such that $(v, \nu(m, v)) \in E$ for all $m \in M$ and $v \in V_i$.

We can extend the memory update function δ to a function $\delta^* : M \times \text{Hist} \rightarrow M$ defined by $\delta^*(m, \epsilon) = m$ and $\delta^*(m, hv) = \delta(\delta^*(m, h), v)$ for all $m \in M$ and $hv \in \text{Hist}$. The strategy $\sigma_{\mathcal{A}_i}$ computed by a finite strategy automaton \mathcal{A}_i is defined by $\sigma_{\mathcal{A}_i}(hv) = \nu(\delta^*(m_0, h), v)$ for all $hv \in \text{Hist}$ such that $v \in V_i$. We say that σ is a *finite-memory strategy* if there exists⁴ a finite strategy automaton \mathcal{A} such that $\sigma = \sigma_{\mathcal{A}}$. Moreover, we say that $\sigma = \sigma_{\mathcal{A}}$ has a memory of size at most $|M|$, where $|M|$ is the number of states of \mathcal{A} . In particular, if $|M| = 1$, we say that σ is a *positional strategy* (the current vertex of the play determines the choice of the next vertex). We call $(\sigma_i)_{i \in \Pi}$ a strategy profile with memory m if for all $i \in \Pi$, the strategy σ_i has a memory of size at most m . A strategy profile $(\sigma_i)_{i \in \Pi}$ is called *positional* or *finite-memory* if each σ_i is a positional or a finite-memory strategy, respectively.

We now define the notion of *Nash equilibria* in this quantitative framework.

Definition 2. Given an initialised multiplayer cost game (\mathcal{G}, v_0) , a strategy profile $(\sigma_i)_{i \in \Pi}$ is a Nash equilibrium in (\mathcal{G}, v_0) if, for every player $j \in \Pi$ and for every strategy σ'_j of player j , we have:

$$\text{Cost}_j(\rho) \leq \text{Cost}_j(\rho')$$

where $\rho = \langle (\sigma_i)_{i \in \Pi} \rangle_{v_0}$ and $\rho' = \langle \sigma'_j, \sigma_{i \in \Pi \setminus \{j\}} \rangle_{v_0}$.

This definition means that, for all $j \in \Pi$, player j has no incentive to deviate from σ_j since he cannot strictly decrease his cost when using σ'_j instead of σ_j . Keeping notations of Definition 2 in mind, a strategy σ'_j such that $\text{Cost}_j(\rho) > \text{Cost}_j(\rho')$ is called a *profitable deviation* for player j w.r.t. $(\sigma_i)_{i \in \Pi}$.

Example 3. Let $\mathcal{G} = (\Pi, V, V_1, V_2, E, \text{Cost}_1, \text{Cost}_2)$ be the two-player cost game whose graph $G = (V, E)$ is depicted in Figure 1. The states of player 1 (resp. 2) are represented by circles (resp. squares)⁵. Thus, according to Figure 1, $V_1 = \{A, C, D\}$ and $V_2 = \{B\}$. In order to define the cost functions of both players, we consider a price function $\pi : E \rightarrow \{1, 2, 3\}$, which assigns a price to each edge of the graph. The price function⁶ π is as follows (see the numbers in Figure 1): $\pi(A, B) = \pi(B, A) = \pi(B, C) = 1$, $\pi(A, D) = 2$ and $\pi(C, B) = \pi(D, B) = 3$. The cost function Cost_1 of player 1 expresses a *quantitative reachability objective*: he wants to reach the vertex C

⁴ Note that there exist several finite strategy automata such that $\sigma = \sigma_{\mathcal{A}}$.

⁵ We will keep this convention through the paper.

⁶ Note that we could have defined a different price function for each player. In this case, the edges of the graph would have been labelled by couples of numbers.

(shaded vertex) while minimising the sum of prices up to this vertex. That is, for every play $\rho = \rho_0 \rho_1 \dots$ of \mathcal{G} :

$$\text{Cost}_1(\rho) = \begin{cases} \sum_{i=1}^n \pi(\rho_{i-1}, \rho_i) & \text{if } n \text{ is the least index s.t. } \rho_n = C, \\ +\infty & \text{otherwise.} \end{cases}$$

As for the cost function Cost_2 of player 2, it expresses a *mean-payoff objective*: the cost of a play is the long-run average of the prices that appear along this play. Formally, for any play $\rho = \rho_0 \rho_1 \dots$ of \mathcal{G} :

$$\text{Cost}_2(\rho) = \limsup_{n \rightarrow +\infty} \frac{1}{n} \cdot \sum_{i=1}^n \pi(\rho_{i-1}, \rho_i).$$

Each player aims at minimising the cost incurred by the play. Let us insist on the fact that the players of a cost game may have different kinds of cost functions (as in this example).

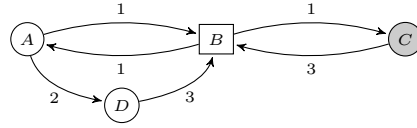


Fig. 1. A two-player cost game \mathcal{G} .

An example of a play in \mathcal{G} can be given by $\rho = (AB)^\omega$, leading to the costs $\text{Cost}_1(\rho) = +\infty$ and $\text{Cost}_2(\rho) = 1$. In the same way, the play $\rho' = A(BC)^\omega$ induces the following costs: $\text{Cost}_1(\rho') = 2$ and $\text{Cost}_2(\rho') = 2$.

Let us fix the initial vertex v_0 at the vertex A . The play $\rho = (AB)^\omega$ is the outcome of the positional strategy⁷ profile (σ_1, σ_2) where $\sigma_1(A) = B$ and $\sigma_2(B) = A$. Moreover, this strategy profile is in fact a *Nash equilibrium*: player 2 gets the least cost he can expect in this game, and player 1 has no incentive to choose the edge (A, D) (it does not allow the play to pass through vertex C).

We now consider the positional strategy profile (σ'_1, σ'_2) with $\sigma'_1(A) = B$ and $\sigma'_2(B) = C$. Its outcome is the play $\rho' = A(BC)^\omega$. However, this strategy profile is *not* a Nash equilibrium, because player 2 can strictly lower his cost by always choosing the edge (B, A) instead of (B, C) , thus lowering his cost from 2 to 1. In other words, the strategy σ_2 (defined before) is a *profitable deviation* for player 2 w.r.t. (σ'_1, σ'_2) .

The questions studied in this paper are the following ones:

Problem 1 Given a multiplayer cost game \mathcal{G} , does there exist a Nash equilibrium in \mathcal{G} ?

Problem 2 Given a multiplayer cost game \mathcal{G} , does there exist a finite-memory Nash equilibrium in \mathcal{G} ?

⁷ Note that player 1 has no choice in vertices C and D , that is, $\sigma_1(hv)$ is necessarily equal to B for $v \in \{C, D\}$ and $h \in \text{Hist}$.

Obviously enough, if we make no restrictions on our cost games, the answer to Problem 1 (and thus to Problem 2) is negative (see Example 4). Our first goal in this paper is to identify a large class of cost games for which the answer to Problem 1 is positive. Then we also positively reply to Problem 2 for subclasses of the previously identified class of cost games. Both results can be found in Section 3.

Example 4. Let (\mathcal{G}, A) be the initialised one-player cost game depicted below, whose cost function Cost_1 is defined by $\text{Cost}_1(A^n B^\omega) = \frac{1}{n}$ for $n \in \mathbb{N}_0$ and $\text{Cost}_1(A^\omega) = +\infty$. One can be convinced that there is no Nash equilibrium in this initialised game.



In order to our class of cost games, we need the notions of *Min-Max cost games*, *determinacy* and *optimal strategies*. The following two definitions are inspired by [27].

Definition 5. A Min-Max cost game is a tuple $\mathcal{G} = (V, V_{\text{Min}}, V_{\text{Max}}, E, \text{Cost}_{\text{Min}}, \text{Gain}_{\text{Max}})$, where

- $G = (V, E)$ is a finite directed graph with vertices V and edges $E \subseteq V \times V$,
- $(V_{\text{Min}}, V_{\text{Max}})$ is a partition of V such that V_{Min} (resp. V_{Max}) is the set of vertices controlled by player Min (resp. Max), and
- $\text{Cost}_{\text{Min}} : \text{Plays} \rightarrow \mathbb{R} \cup \{+\infty, -\infty\}$ is the cost function of player Min, that represents the amount that he loses for a play, and $\text{Gain}_{\text{Max}} : \text{Plays} \rightarrow \mathbb{R} \cup \{+\infty, -\infty\}$ is the gain function of player Max, that represents the amount that he wins for a play.

In such a game, player Min wants to *minimise* his cost, while player Max wants to *maximise* his gain. So, a Min-Max cost game is a particular case of a two-player cost game. Let us stress that, according to this definition, a Min-Max cost game is *zero-sum* if $\text{Cost}_{\text{Min}} = \text{Gain}_{\text{Max}}$, but this might not always be the case⁸. We also point out that Definition 5 allows to take completely unrelated functions Cost_{Min} and Gain_{Max} , but usually they are similar (see Definition 15). In the sequel, we denote by Σ_{Min} (resp. Σ_{Max}) the set of strategies of player Min (resp. Max) in a Min-Max cost game.

Definition 6. Given a Min-Max cost game \mathcal{G} , we define for every vertex $v \in V$ the upper value $\text{Val}^*(v)$ as:

$$\text{Val}^*(v) = \inf_{\sigma_1 \in \Sigma_{\text{Min}}} \sup_{\sigma_2 \in \Sigma_{\text{Max}}} \text{Cost}_{\text{Min}}(\langle \sigma_1, \sigma_2 \rangle_v),$$

and the lower value $\text{Val}_*(v)$ as:

$$\text{Val}_*(v) = \sup_{\sigma_2 \in \Sigma_{\text{Max}}} \inf_{\sigma_1 \in \Sigma_{\text{Min}}} \text{Gain}_{\text{Max}}(\langle \sigma_1, \sigma_2 \rangle_v).$$

The game \mathcal{G} is determined if, for every $v \in V$, we have $\text{Val}^*(v) = \text{Val}_*(v)$. In this case, we say that the game \mathcal{G} has a value, and for every $v \in V$, $\text{Val}(v) = \text{Val}^*(v) =$

⁸ For an example, see the average-price game in Definition 15.

$\text{Val}_*(v)$. We also say that the strategies $\sigma_1^* \in \Sigma_{\text{Min}}$ and $\sigma_2^* \in \Sigma_{\text{Max}}$ are optimal strategies for the respective players if, for every $v \in V$, we have that

$$\inf_{\sigma_1 \in \Sigma_{\text{Min}}} \text{Gain}_{\text{Max}}(\langle \sigma_1, \sigma_2^* \rangle_v) = \text{Val}(v) = \sup_{\sigma_2 \in \Sigma_{\text{Max}}} \text{Cost}_{\text{Min}}(\langle \sigma_1^*, \sigma_2 \rangle_v).$$

If σ_1^* is an optimal strategy for player Min, then he loses at most $\text{Val}(v)$ when playing according to it. On the other hand, player Max wins at least $\text{Val}(v)$ if he plays according to an optimal strategy σ_2^* for him.

Examples of classical determined Min-Max cost games can be found in Section 4.

3 Results

In this section, we first define a large class of cost games for which Problem 1 can be answered positively (Theorem 10). Then, we study existence of simple Nash equilibria (Theorems 13 and 14). To define this interesting class of cost games, we need the concepts of *cost-prefix-linear* and *coalition-determined* cost games.

Definition 7. A multiplayer cost game $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, E, (\text{Cost}_i)_{i \in \Pi})$ is cost-prefix-linear if, for every player $i \in \Pi$, every vertex $v \in V$ and history $hv \in \text{Hist}$, there exists $a \in \mathbb{R}$ and $b \in \mathbb{R}^+$ such that, for every play $\rho \in \text{Plays}$ with $\text{First}(\rho) = v$, we have:

$$\text{Cost}_i(h\rho) = a + b \cdot \text{Cost}_i(\rho).$$

Let us now define the concept of *coalition-determined* cost games.

Definition 8. A multiplayer cost game $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, E, (\text{Cost}_i)_{i \in \Pi})$ is (positionally/finite-memory) coalition-determined if, for every player $i \in \Pi$, there exists a gain function $\text{Gain}_{\text{Max}}^i : \text{Plays} \rightarrow \mathbb{R} \cup \{+\infty, -\infty\}$ such that

- $\text{Cost}_i \geq \text{Gain}_{\text{Max}}^i$, and
- the Min-Max cost game $\mathcal{G}^i = (V, V_i, V \setminus V_i, E, \text{Cost}_i, \text{Gain}_{\text{Max}}^i)$, where player i (player Min) plays against the coalition $\Pi \setminus \{i\}$ (player Max), is determined and has (positional/finite-memory) optimal strategies for both players. That is: $\exists \sigma_i^* \in \Sigma_{\text{Min}}, \exists \sigma_{-i}^* \in \Sigma_{\text{Max}}$ (both positional/finite-memory) such that $\forall v \in V$

$$\inf_{\sigma_i \in \Sigma_{\text{Min}}} \text{Gain}_{\text{Max}}^i(\langle \sigma_i, \sigma_{-i}^* \rangle_v) = \text{Val}^i(v) = \sup_{\sigma_{-i} \in \Sigma_{\text{Max}}} \text{Cost}_i(\langle \sigma_i^*, \sigma_{-i} \rangle_v).$$

Given $i \in \Pi$, note that \mathcal{G}^i does not depend on the cost functions Cost_j , with $j \neq i$.

Example 9. Let us consider the two-player cost game \mathcal{G} of Example 3, where player 1 has a quantitative reachability objective (Cost_1) and player 2 has a mean-payoff objective (Cost_2). We show that \mathcal{G} is positionally coalition-determined.

Let us set $\text{Gain}_{\text{Max}}^1 = \text{Cost}_1$ and study the Min-Max cost game $\mathcal{G}^1 = (V, V_1, V_2, E, \text{Cost}_1, \text{Gain}_{\text{Max}}^1)$, where player Min (resp. Max) is player 1 (resp. 2) and wants to minimise Cost_1 (resp. maximise $\text{Gain}_{\text{Max}}^1$). This game is positionally determined [27, 14]. We define positional strategies σ_1^* and σ_{-1}^* for player 1 and player 2, respectively,

in the following way: $\sigma_1^*(A) = B$ and $\sigma_{-1}^*(B) = A$. From A , their outcome is $\langle(\sigma_1^*, \sigma_{-1}^*)\rangle_A = (AB)^\omega$, and $\text{Cost}_1((AB)^\omega) = \text{Gain}_{\text{Max}}^1((AB)^\omega) = +\infty$. One can check that the strategies σ_1^* and σ_{-1}^* are optimal in \mathcal{G}^1 . Note that the positional strategy $\tilde{\sigma}_1^*$ defined by $\tilde{\sigma}_1^*(A) = D$ is also optimal (for player 1) in \mathcal{G}^1 . With this strategy, we have that $\langle(\tilde{\sigma}_1^*, \sigma_{-1}^*)\rangle_A = (ADB)^\omega$, and $\text{Cost}_1((ADB)^\omega) = \text{Gain}_{\text{Max}}^1((ADB)^\omega) = +\infty$.

We now examine the Min-Max cost game $\mathcal{G}^2 = (V, V_2, V_1, E, \text{Cost}_2, \text{Gain}_{\text{Max}}^2)$, where $\text{Gain}_{\text{Max}}^2$ is defined as Cost_2 but with \liminf instead of \limsup . In this game, player Min (resp. Max) is player 2 (resp. 1) and wants to minimise Cost_2 (resp. maximise $\text{Gain}_{\text{Max}}^2$). This game is also positionally determined [27,14]. Let σ_2^* and σ_{-2}^* be the positional strategies for player 2 and player 1, respectively, defined as follows: $\sigma_2^*(B) = C$ and $\sigma_{-2}^*(A) = D$. From A , their outcome is $\langle(\sigma_2^*, \sigma_{-2}^*)\rangle_A = AD(BC)^\omega$, and $\text{Cost}_2(AD(BC)^\omega) = \text{Gain}_{\text{Max}}^2(AD(BC)^\omega) = 2$. We claim that σ_2^* and σ_{-2}^* are the only positional optimal strategies in \mathcal{G}^2 .

Theorem 10 positively answers Problem 1 for cost-prefix-linear, coalition-determined cost games.

Theorem 10. *In every initialised multiplayer cost game that is cost-prefix-linear and coalition-determined, there exists a Nash equilibrium.*

Proof. Let $(\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, E, (\text{Cost}_i)_{i \in \Pi}), v_0)$ be an initialised multiplayer cost game that is cost-prefix-linear and coalition-determined. Thanks to the latter property, we know that, for every $i \in \Pi$, there exists a gain function $\text{Gain}_{\text{Max}}^i$ such that the Min-Max cost game $\mathcal{G}^i = (V, V_i, V \setminus V_i, E, \text{Cost}_i, \text{Gain}_{\text{Max}}^i)$ is determined and there exist optimal strategies σ_i^* and σ_{-i}^* for player i and the coalition $\Pi \setminus \{i\}$ respectively. In particular, for $j \neq i$, we denote by $\sigma_{j,i}^*$ the strategy of player j derived from the strategy σ_{-i}^* of the coalition $\Pi \setminus \{i\}$.

The idea is to define the required Nash equilibrium as follows: each player i plays according to his strategy σ_i^* and punishes the first player $j \neq i$ who deviates from his strategy σ_j^* , by playing according to $\sigma_{i,j}^*$ (the strategy of player i derived from σ_{-j}^* in the game \mathcal{G}^j).

Formally, we consider the outcome of the optimal strategies $(\sigma_i^*)_{i \in \Pi}$ from v_0 , and set $\rho := \langle(\sigma_i^*)_{i \in \Pi}\rangle_{v_0}$. We need to specify a punishment function $P : \text{Hist} \rightarrow \Pi \cup \{\perp\}$ that detects who is the first player to deviate from the play ρ , i.e. who has to be punished. For the initial vertex v_0 , we define $P(v_0) = \perp$ (meaning that nobody has deviated from ρ) and for every history $hv \in \text{Hist}$, we let:

$$P(hv) := \begin{cases} \perp & \text{if } P(h) = \perp \text{ and } hv \text{ is a prefix of } \rho, \\ i & \text{if } P(h) = \perp, hv \text{ is not a prefix of } \rho, \text{ and } \text{Last}(h) \in V_i, \\ P(h) & \text{otherwise } (P(h) \neq \perp). \end{cases}$$

Then the definition of the Nash equilibrium $(\tau_i)_{i \in \Pi}$ in \mathcal{G} is as follows. For all $i \in \Pi$ and $h \in \text{Hist}$ such that $\text{Last}(h) \in V_i$,

$$\tau_i(h) := \begin{cases} \sigma_i^*(h) & \text{if } P(h) = \perp \text{ or } i, \\ \sigma_{i,P(h)}^*(h) & \text{otherwise.} \end{cases}$$

Clearly the outcome of $(\tau_i)_{i \in \Pi}$ is the play $\rho (= \langle(\sigma_i^*)_{i \in \Pi}\rangle_{v_0})$.

Now we show that the strategy profile $(\tau_i)_{i \in \Pi}$ is a Nash equilibrium in \mathcal{G} . As a contradiction, let us assume that there exists a profitable deviation τ'_j for some player $j \in \Pi$. We denote by $\rho' := \langle \tau'_j, (\tau_i)_{i \in \Pi \setminus \{j\}} \rangle_{v_0}$ the outcome where player j plays according to his profitable deviation τ'_j and the players of the coalition $\Pi \setminus \{j\}$ keep their strategies $(\tau_i)_{i \in \Pi \setminus \{j\}}$. Since τ'_j is a profitable deviation for player j w.r.t. $(\tau_i)_{i \in \Pi}$, we have that:

$$\text{Cost}_j(\rho') < \text{Cost}_j(\rho). \quad (1)$$

As both plays ρ and ρ' start from vertex v_0 , there exists a history $hv \in \text{Hist}$ such that $\rho = h \langle (\tau_i)_{i \in \Pi} \rangle_v$ and $\rho' = h \langle \tau'_j, (\tau_i)_{i \in \Pi \setminus \{j\}} \rangle_v$ (remark that h could be empty). Among the common prefixes of ρ and ρ' , we choose the history hv of maximal length. By definition of the strategy profile $(\tau_i)_{i \in \Pi}$, we can write in the case of the outcome ρ that $\rho = h \langle (\sigma_i^*)_{i \in \Pi} \rangle_v$. Whereas in the case of the outcome ρ' , player j does not follow his strategy σ_j^* any more from vertex v , and so, the coalition $\Pi \setminus \{j\}$ punishes him by playing according to the strategy σ_{-j}^* after history hv , and so $\rho' = h \langle \tau'_j, \sigma_{-j}^* \rangle_v$ (see Figure 2).

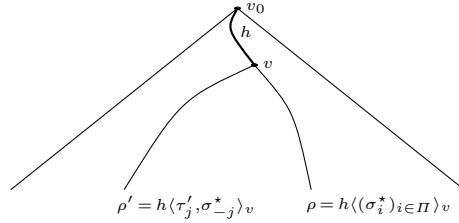


Fig. 2. Sketch of the tree representing the unravelling of the game \mathcal{G} from v_0 .

Since σ_{-j}^* is an optimal strategy for the coalition $\Pi \setminus \{j\}$ in the determined Min-Max cost game \mathcal{G}^j , we have:

$$\begin{aligned} \text{Val}^j(v) &= \inf_{\sigma_j \in \Sigma_{\text{Min}}} \text{Gain}_{\text{Max}}^j(\langle \sigma_j, \sigma_{-j}^* \rangle_v) \\ &\leq \text{Gain}_{\text{Max}}^j(\langle \tau'_j, \sigma_{-j}^* \rangle_v) \\ &\leq \text{Cost}_j(\langle \tau'_j, \sigma_{-j}^* \rangle_v). \end{aligned} \quad (2)$$

The last inequality comes from the hypothesis $\text{Cost}_j \geq \text{Gain}_{\text{Max}}^j$ in the game \mathcal{G}^j .

Moreover, the game \mathcal{G} is cost-prefix-linear, and then, when considering the history hv , there exist $a \in \mathbb{R}$ and $b \in \mathbb{R}^+$ such that

$$\text{Cost}_j(\rho') = \text{Cost}_j(h \langle \tau'_j, \sigma_{-j}^* \rangle_v) = a + b \cdot \text{Cost}_j(\langle \tau'_j, \sigma_{-j}^* \rangle_v). \quad (3)$$

As $b \geq 0$, Equations (2) and (3) imply:

$$\text{Cost}_j(\rho') \geq a + b \cdot \text{Val}^j(v). \quad (4)$$

Since h is also a prefix of ρ , we have:

$$\text{Cost}_j(\rho) = \text{Cost}_j(h \langle (\sigma_i^*)_{i \in \Pi} \rangle_v) = a + b \cdot \text{Cost}_j(\langle (\sigma_i^*)_{i \in \Pi} \rangle_v). \quad (5)$$

Furthermore, as σ_j^* is an optimal strategy for player j in the Min-Max cost game \mathcal{G}^j , it follows that:

$$\begin{aligned} \text{Val}^j(v) &= \sup_{\sigma_{-j} \in \Sigma_{\text{Max}}} \text{Cost}_j(\langle \sigma_j^*, \sigma_{-j} \rangle_v) \\ &\geq \text{Cost}_j(\langle (\sigma_i^*)_{i \in \Pi} \rangle_v). \end{aligned} \quad (6)$$

Then, Equations (5) and (6) imply:

$$\text{Cost}_j(\rho) \leq a + b \cdot \text{Val}^j(v). \quad (7)$$

Finally, Equations (4) and (7) lead to the following inequality:

$$\text{Cost}_j(\rho) \leq a + b \cdot \text{Val}^j(v) \leq \text{Cost}_j(\rho'),$$

which contradicts Equation (1). The strategy profile $(\tau_i)_{i \in \Pi}$ is then a Nash equilibrium in the game \mathcal{G} . \square

Remark 11. The proof of Theorem 10 remains valid for cost functions $\text{Cost}_i : \text{Plays} \rightarrow K$, where $\langle K, +, \cdot, 0, 1, \leq \rangle$ is an ordered field. This allows for instance to consider non-standard real costs and enjoy infinitesimals to model the costs of a player.

Example 12. Let us consider the initialised two-player cost game (\mathcal{G}, A) of Example 3, where player 1 has a quantitative reachability objective (Cost_1) and player 2 has a mean-payoff objective (Cost_2). One can show that \mathcal{G} is cost-prefix-linear. Since we saw in Example 9 that this game is also positionally coalition-determined, we can apply the construction in the proof of Theorem 10 to get a Nash equilibrium in \mathcal{G} . The construction from this proof may result in two different Nash equilibria, depending on the selection of the strategies $\sigma_1^*/\tilde{\sigma}_1^*$, σ_{-1}^* , σ_2^* and σ_{-2}^* as defined in Example 9.

The first Nash equilibrium (τ_1, τ_2) with outcome $\rho = \langle \sigma_1^*, \sigma_2^* \rangle_A = A(BC)^\omega$ is given, for any history h , by:

$$\tau_1(hA) := \begin{cases} B & \text{if } P(hA) = \{\perp, 1\} \\ D & \text{otherwise} \end{cases} \quad ; \quad \tau_2(hB) := \begin{cases} C & \text{if } P(hB) = \{\perp, 2\} \\ A & \text{otherwise} \end{cases}$$

where the punishment function P is defined as in the proof of Theorem 10 and depends on the play ρ . The cost for this finite-memory Nash equilibrium is $\text{Cost}_1(\rho) = 2 = \text{Cost}_2(\rho)$.

The strategy $\tilde{\tau}_1$ of the second Nash equilibrium $(\tilde{\tau}_1, \tau_2)$ with outcome $\tilde{\rho} = \langle \tilde{\sigma}_1^*, \sigma_2^* \rangle_A = AD(BC)^\omega$ is given by $\tilde{\tau}_1(hA) := D$ for all history h . The cost for this finite-memory Nash equilibrium is $\text{Cost}_1(\tilde{\rho}) = 6$ and $\text{Cost}_2(\tilde{\rho}) = 2$, respectively.

Note that there is no positional Nash equilibrium with outcome ρ (resp. $\tilde{\rho}$).

The two following theorems provide results about the complexity of the Nash equilibrium defined in the latter proof. Applications of these theorems to specific classes of cost games are provided in Section 4.

Theorem 13. *In every initialised multiplayer cost game that is cost-prefix-linear and positionally coalition-determined, there exists a Nash equilibrium with memory (at most) $|V| + |\Pi|$.*

Theorem 14. *In every initialised multiplayer cost game that is cost-prefix-linear and finite-memory coalition-determined, there exists a Nash equilibrium with finite memory.*

The proofs of these two theorems rely on the construction of the Nash equilibrium provided in the proof of Theorem 10.

4 Applications

In this section, we exhibit several classes of *classical objectives* that can be encoded in our general setting. The list we propose is far from being exhaustive.

4.1 Qualitative Objectives

Multiplayer games with qualitative (win/lose) objectives can naturally be encoded via multiplayer cost games; for instance via cost functions $\text{Cost}_i : \text{Plays} \rightarrow \{1, +\infty\}$, where 1 (resp. $+\infty$) means that the play is won (resp. lost) by player i . Let us now consider the subclass of qualitative games with prefix-independent⁹ Borel objectives. Given such a game \mathcal{G} , we have that \mathcal{G} is coalition-determined, as a consequence of the Borel determinacy theorem [20]. Moreover the prefix-independence hypothesis obviously guarantees that \mathcal{G} is also cost-prefix-linear (by taking $a = 0$ and $b = 1$). By applying Theorem 10, we obtain the existence of a Nash equilibrium for qualitative games with prefix-independent Borel objectives. Let us notice that this result is already present in [16].

When considering more specific subclasses of qualitative games enjoying a positional determinacy result, such as parity games [15], we can apply Theorem 13 and ensure existence of a Nash equilibrium whose memory is (at most) linear.

4.2 Classical Quantitative Objectives

We here give four well-known kinds of Min-Max cost games and see later that they are determined. For each sort of game, the cost and gain functions are defined from a price function (and a reward function in the last case), which labels the edges of the game graph with prices (and rewards).

Definition 15 ([27]). *Given a game graph $G = (V, V_{\text{Min}}, V_{\text{Max}}, E)$, a price function $\pi : E \rightarrow \mathbb{R}$ that assigns a price to each edge, a diverging¹⁰ reward function $\vartheta : E \rightarrow \mathbb{R}$ that assigns a reward to each edge, and a play $\rho = \rho_0\rho_1 \dots$ in G , we define the following Min-Max cost games:*

⁹ An objective $\Omega \subseteq V^\omega$ is prefix-independent if only if for every play $\rho = \rho_0\rho_1 \dots \in V^\omega$, we have that $\rho \in \Omega$ iff for every $n \in \mathbb{N}$, $\rho_n\rho_{n+1} \dots \in \Omega$.

¹⁰ For all plays $\rho = \rho_0\rho_1 \dots$ in G , it holds that $\lim_{n \rightarrow \infty} |\sum_{i=1}^n \vartheta(\rho_{i-1}, \rho_i)| = +\infty$. This is equivalent to requiring that every cycle has a positive sum of rewards.

- (i) a reachability-price game is a Min-Max cost game $\mathcal{G} = (G, RP_{Min}, RP_{Max})$ together with a given goal set $Goal \subseteq V$, where

$$RP_{Min}(\rho) = RP_{Max}(\rho) = \begin{cases} \pi(\rho[0, n]) & \text{if } n \text{ is the least index s.t. } \rho_n \in Goal, \\ +\infty & \text{otherwise,} \end{cases}$$

with $\pi(\rho[0, n]) = \sum_{i=1}^n \pi(\rho_{i-1}, \rho_i)$;

- (ii) a discounted-price game is a Min-Max cost game $\mathcal{G} = (G, DP_{Min}(\lambda), DP_{Max}(\lambda))$ together with a given discount factor $\lambda \in]0, 1[$, where

$$DP_{Min}(\lambda)(\rho) = DP_{Max}(\lambda)(\rho) = (1 - \lambda) \cdot \sum_{i=1}^{+\infty} \lambda^{i-1} \pi(\rho_{i-1}, \rho_i);$$

- (iii) an average-price game¹¹ is a Min-Max cost game $\mathcal{G} = (G, AP_{Min}, AP_{Max})$, where

$$AP_{Min}(\rho) = \limsup_{n \rightarrow +\infty} \frac{\pi(\rho[0, n])}{n} \quad \text{and} \quad AP_{Max}(\rho) = \liminf_{n \rightarrow +\infty} \frac{\pi(\rho[0, n])}{n};$$

- (iv) a price-per-reward-average game is a Min-Max cost game $\mathcal{G} = (G, PRAvg_{Min}, PRAvg_{Max})$, where

$$PRAvg_{Min}(\rho) = \limsup_{n \rightarrow +\infty} \frac{\pi(\rho[0, n])}{\vartheta(\rho[0, n])} \quad \text{and} \quad PRAvg_{Max}(\rho) = \liminf_{n \rightarrow +\infty} \frac{\pi(\rho[0, n])}{\vartheta(\rho[0, n])},$$

with $\vartheta(\rho[0, n]) = \sum_{i=1}^n \vartheta(\rho_{i-1}, \rho_i)$.

An average-price game is then a particular case of a price-per-reward-average game. Let us remark that, in Example 3, the cost function $Cost_1$ (resp. $Cost_2$) corresponds to RP_{Min} with $Goal = \{C\}$ (resp. AP_{Min}). The game \mathcal{G}^1 (resp. \mathcal{G}^2) of Example 9 is a reachability-price (resp. average-price) game.

The following theorem is a well-known result about the particular cost games described in Definition 15.

Theorem 16 ([27,14]). *Reachability-price games, discounted-price games, average-price games, and price-per-reward games are determined and have positional optimal strategies.*

This result implies that a multiplayer cost game where each cost function is RP_{Min} , DP_{Min} , AP_{Min} or $PRAvg_{Min}$ is positionally coalition-determined. Moreover, one can show that such a game is cost-prefix-linear. Theorem 17 then follows from Theorem 13.

Theorem 17. *In every initialised multiplayer cost game $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, E, (Cost_i)_{i \in \Pi})$ where the cost function $Cost_i$ belongs to $\{RP_{Min}, DP_{Min}, AP_{Min}, PRAvg_{Min}\}$ for every player $i \in \Pi$, there exists a Nash equilibrium with memory (at most) $|V| + |\Pi|$.*

¹¹ When the cost function of a player is AP_{Min} , we say that he has a mean-payoff objective.

Note that the existence of finite-memory Nash equilibria in cost games with quantitative reachability objectives has already been established in [7,8]. Even if not explicitly stated in the previous papers, one can deduce from the proof of [8, Lemma 16] that the provided Nash equilibrium has a memory (at least) exponential in the size of the cost game. Thus, Theorem 17 significantly improves the complexity of the strategies constructed in the case of cost games with quantitative reachability objectives.

4.3 Combining Qualitative and Quantitative Objectives

Multipayer cost games allow to encode games combining both qualitative and quantitative objectives, such as *mean-payoff parity games* [11]. In our framework, where each player aims at minimising his cost, the mean-payoff parity objective could be encoded as follows: $\text{Cost}_i(\rho) = \text{AP}_{\text{Min}}(\rho)$ if the parity condition is satisfied, $+\infty$ otherwise.

The determinacy of mean-payoff parity games, together with the existence of optimal strategies (that could require infinite memory) have been proved in [11]. This result implies that multipayer cost games with mean-payoff parity objectives are coalition-determined. Moreover, one can prove that such a game is also cost-prefix-linear (by taking $a = 0$ and $b = 1$). By applying Theorem 10, we obtain the existence of a Nash equilibrium for multipayer cost games with mean-payoff parity objectives. As far as we know, this is the first result about the existence of a Nash equilibrium in cost games with mean-payoff parity games.

Remark 18. Let us emphasise that Theorem 10 applies to cost games where the players have different kinds of cost functions (as in Example 3). In particular, one player could have a qualitative Büchi objective, a second player a discounted-price objective, a third player a mean-payoff parity objective,...

References

1. R. Alur, A. Kanade, and G. Weiss. Ranking automata and games for prioritized requirements. In *CAV*, volume 5123 of *LNCS*, pages 240–253. Springer, 2008.
2. R. Bloem, K. Chatterjee, T. Henzinger, and B. Jobstmann. Better quality in synthesis through quantitative objectives. In *CAV*, volume 5643 of *LNCS*, pages 140–156. Springer, 2009.
3. E. Boros and V. Gurvich. Why chess and back gammon can be solved in pure positional uniformly optimal strategies. Rutcor Research Report 21-2009, Rutgers University, 2009.
4. P. Bouyer, R. Brenguier, and N. Markey. Nash equilibria for reachability objectives in multipayer timed games. In *Concurrency Theory, CONCUR*, volume 6269 of *LNCS*, pages 192–206. Springer, 2010.
5. P. Bouyer, R. Brenguier, N. Markey, and M. Ummels. Nash equilibria in concurrent games with Büchi objectives. In *FSTTCS*, volume 13 of *LIPICs*, pages 375–386. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2011.
6. P. Bouyer, U. Fahrenberg, K. G. Larsen, N. Markey, and J. Srba. Infinite runs in weighted timed automata with energy constraints. In *FORMATS*, volume 5215 of *LNCS*, pages 33–47. Springer, 2008.
7. T. Brihaye, V. Bruyère, and J. De Pril. Equilibria in quantitative reachability games. In *CSR*, volume 6072 of *LNCS*, pages 72–83. Springer, 2010.

8. T. Brihaye, V. Bruyère, and J. De Pril. On equilibria in quantitative games with reachability/safety objectives. *CoRR*, abs/1205.4889, 2012.
9. T. Brihaye, V. Bruyère, J. De Pril, and H. Gimbert. Subgame perfection for equilibria in quantitative reachability games. In *FoSSaCS*, volume 7213 of *LNCS*, pages 286–300. Springer, 2012.
10. K. Chatterjee and T. A. Henzinger. Finitary winning in omega-regular games. In *TACAS*, volume 3920 of *LNCS*, pages 257–271. Springer, 2006.
11. K. Chatterjee, T. A. Henzinger, and M. Jurdzinski. Mean-payoff parity games. In *LICS*, pages 178–187. IEEE Computer Society, 2005.
12. E. Clarke, O. Grumberg, and D. Peled. *Model Checking*. MIT Press, Cambridge, MA, 2000.
13. N. Fijalkow and M. Zimmermann. Cost-parity and cost-streett games. *CoRR*, abs/1207.0663, 2012.
14. J. Filar and K. Vrieze. *Competitive Markov decision processes*. Springer Verlag, 1997.
15. E. Grädel, W. Thomas, and T. Wilke, editors. *Automata, logics, and infinite games*, volume 2500 of *LNCS*. Springer, 2002.
16. E. Grädel and M. Ummels. Solution concepts and algorithms for infinite multiplayer games. In *New Perspectives on Games and Interaction*, volume 4 of *Texts in Logic and Games*, pages 151–178. Amsterdam University Press, 2008.
17. J. E. Hopcroft and J. D. Ullman. *Introduction to automata theory, languages, and computation*. Addison-Wesley Publishing Co., Reading, Mass., 1979. Addison-Wesley Series in Computer Science.
18. F. Horn, W. Thomas, and N. Wallmeier. Optimal strategy synthesis in request-response games. In *ATVA*, volume 5311 of *LNCS*, pages 361–373. Springer, 2008.
19. M. Klimoš, K. Larsen, F. Štefaňák, and J. Thaarup. Nash Equilibria in Concurrent Priced Games. In *LATA*, volume 7183 of *LNCS*, pages 363–376. Springer, 2012.
20. D. A. Martin. Borel determinacy. *Ann. of Math. (2)*, 102(2):363–371, 1975.
21. J. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences of the United States of America*, 36(1):48–49, 1950.
22. M. Osborne and A. Rubinstein. *A course in game theory*. MIT Press, Cambridge, MA, 1994.
23. S. Paul, S. Simon, R. Kannan, and K. Kumar. Nash equilibrium in generalised muller games. In *FSTTCS*, volume 4 of *LIPICs*, pages 335–346. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2009.
24. W. Thomas. On the synthesis of strategies in infinite games. In *STACS*, volume 900 of *LNCS*, pages 1–13. Springer, Berlin, 1995.
25. W. Thomas. Church’s problem and a tour through automata theory. In *Pillars of Computer Science*, volume 4800 of *LNCS*, pages 635–655. Springer, 2008.
26. F. Thuijsman and T. E. S. Raghavan. Perfect information stochastic games and related classes. *International Journal of Game Theory*, 26(3):403–408, 1998.
27. A. Trivedi. *Competative optimisation on timed automata*. PhD thesis, University of Warwick, 2009.
28. M. Zimmermann. Time-optimal winning strategies for poset games. In *CIAA*, volume 5642 of *LNCS*, pages 217–226. Springer, 2009.

Technical Appendix

A Example of a cost game which is not cost-prefix-linear

Example 19. Multiplayer cost games allow to encode energy games. Let \mathcal{G} be a cost game defined by means of a price function $\pi : E \rightarrow \mathbb{R}$, that assigns a price to each edge. In our framework, where each player aims at minimising his cost, an energy objective [6] (with threshold $T \in \mathbb{R}$) could be encoded as follows:

$$\text{Cost}_i(\rho) = \begin{cases} \sup_{n \geq 0} \pi(\rho[0, n]) & \text{if } \sup_{n \geq 0} \pi(\rho[0, n]) \leq T \\ +\infty & \text{otherwise,} \end{cases}$$

with $\pi(\rho[0, n]) = \sum_{i=1}^n \pi(\rho_{i-1}, \rho_i)$.

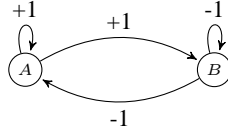


Fig. 3. A cost game which is not cost-prefix-linear

Let us consider the one-player cost game with an energy objective (with threshold $T = 2$) depicted in Figure 3. We show that this game is not cost-prefix-linear. For this, we exhibit a history $hv \in \text{Hist}$ such that for all $a, b \in \mathbb{R}$ there exists a play $\rho \in \text{Plays}$ with $\text{First}(\rho) = v$, such that $\text{Cost}_1(h\rho) \neq a + b \cdot \text{Cost}_1(\rho)$. We in fact give a play ρ independent of a and b . Let hv be the history $AAABA$ and ρ be the play $(AB)^\omega$. We have that $\text{Cost}_1(\rho) = 1$ and $\text{Cost}_1(h\rho) = \text{Cost}_1(AA(AB)^\omega) = +\infty$, since $\sup_{n \geq 0} \pi((h\rho)[0, n]) = 3$, which is above the threshold $T = 2$. It is thus impossible to find $a, b \in \mathbb{R}$ such that:

$$+\infty = \text{Cost}_1(h\rho) = a + b \cdot \text{Cost}_1(\rho) = a + b.$$

B Remark about secure and subgame perfect equilibria

Remark 20. It would be tempting to try to prove the existence of *subgame perfect equilibria* or *secure equilibria*¹² in multiplayer cost games with techniques similar to the proof of Theorem 10. However, our definition of the Nash equilibrium in the proof of Theorem 10 is (in general) neither a subgame perfect equilibrium, nor a secure equilibrium. To see this, let us consider the following two cost games \mathcal{G} and \mathcal{H} , whose graphs are depicted on Figure 4 and 5 respectively. Both games are initialised in vertex A .

The game \mathcal{G} is a two-player cost game where the vertices of player 1 (resp. 2) are represented by circles (resp. squares), that is, $V_1 = \{B, C, D, E, F\}$ and $V_2 = \{A\}$.

¹² The definitions of subgame perfect and secure equilibria in this context can be found in [9].

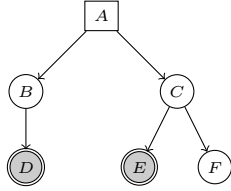


Fig. 4. Game \mathcal{G} .

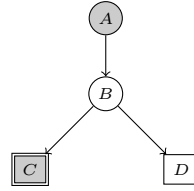


Fig. 5. Game \mathcal{H} .

The cost functions of both players are RP_{Min} , with¹³ $\text{Goal}_1 = \text{Goal}_2 = \{D, E\}$ and the price function $\pi : E \rightarrow \mathbb{R}$ defined by $\pi(e) = 1$ for any edge $e \in E$ (same price function for the two players). It means that both players have reachability objectives and want to reach vertex D or E within the least number of edges.

Let us study the two Min-Max cost games \mathcal{G}^1 and \mathcal{G}^2 . In the game \mathcal{G}^1 , let σ_1^* be defined as $\sigma_1^*(C) = E$ and σ_{-1}^* be defined as $\sigma_{-1}^*(A) = C$. Then, σ_1^* and σ_{-1}^* are positional optimal strategies for player Min (player 1) and player Max (player 2) respectively. In the game \mathcal{G}^2 , we define σ_2^* and σ_{-2}^* as $\sigma_2^*(A) = B$ and $\sigma_{-2}^*(C) = F$. These two strategies of \mathcal{G}^2 are positional optimal strategies for player Min (player 2) and player Max (player 1) respectively.

If we define a Nash equilibrium (τ_1, τ_2) in \mathcal{G} exactly as in the proof of Theorem 10, depending on these strategies σ_1^* , σ_{-1}^* , σ_2^* and σ_{-2}^* , then (τ_1, τ_2) is not a subgame perfect equilibrium in \mathcal{G} . Indeed, $(\tau_1|_A, \tau_2|_A)$ is not a Nash equilibrium in the subgame $\mathcal{G}|_A$ with history AC : player 1 punishes player 2 by choosing the edge (C, F) (according to σ_{-2}^*) whereas player 1 could pay a smaller cost by choosing the edge (C, E) .

Furthermore, this Nash equilibrium also gives a counter-example of subgame perfect equilibrium for other classical punishments (see [22], e.g., punish the last player who has deviated and only for a finite number of steps).

Let us now consider the two-player cost game \mathcal{H} where $V_1 = \{A, B\}$ and $V_2 = \{C, D\}$ (see Figure 5). The price function and the cost functions of the two players are the same as in the game \mathcal{G} , except that here $\text{Goal}_1 = \{A, C\}$ and $\text{Goal}_2 = \{C\}$. Note that player 2 does not really play in \mathcal{H} , only player 1 has a choice to make: he can choose the edge (B, C) or the edge (B, D) .

As before, we study the two Min-Max cost games \mathcal{H}^1 and \mathcal{H}^2 . Let σ_1^* be a positional strategy of player 1 in \mathcal{H}^1 such that $\sigma_1^*(B) = C$, and σ_{-2}^* be a positional strategy of player 1 in \mathcal{H}^2 such that $\sigma_{-2}^*(B) = D$. These strategies are optimal in the two respective games. Then, we define a Nash equilibrium in \mathcal{H} in the same way as in the proof of Theorem 10, depending on σ_1^* and σ_{-2}^* . Actually, this is not a secure equilibrium in \mathcal{H} because player 1 can strictly increase player 2's cost while keeping his own cost, by choosing the edge (B, D) instead of following σ_1^* (σ_1^* suggests to choose the edge (B, C)).

¹³ In both figures, shaded (resp. doubly circled) vertices represent the goal set Goal_1 (resp. Goal_2).

C Proof of Theorem 13

Theorem 13 states that in every initialised multiplayer cost game that is cost-prefix-linear and *positionally* coalition-determined, there exists a Nash equilibrium with memory (at most) $|V| + |\Pi|$.

Proof. Let $(\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, E, (\text{Cost}_i)_{i \in \Pi}), v_0)$ be an initialised multiplayer cost game that is cost-prefix-linear and *positionally* coalition-determined. For this proof, we keep the notations introduced in the proof of Theorem 10. In particular, we consider the Nash equilibrium $(\tau_i)_{i \in \Pi}$ as defined in the latter proof, whose outcome is $\rho := \langle (\sigma_i^*)_{i \in \Pi} \rangle_{v_0}$. We recall that for all $i \in \Pi$, the strategy τ_i depends on the strategies σ_i^* (optimal strategy in \mathcal{G}^i) and $\sigma_{i,j}^*$ (derived from the optimal strategy σ_{-j}^* in \mathcal{G}^j) for $j \in \Pi \setminus \{i\}$. As the game \mathcal{G} is now positionally coalition-determined by hypothesis, these strategies are assumed to be positional. This proof consists in showing that $(\tau_i)_{i \in \Pi}$ is a strategy profile with memory (at most) $|V| + |\Pi|$.

For this purpose, we define a finite strategy automaton for each player that remembers the play ρ and who has to be punished. As the play ρ is the outcome of the positional strategy profile $(\sigma_i^*)_{i \in \Pi}$, we can write $\rho := v_0 \dots v_{k-1} (v_k \dots v_n)^\omega$ where $0 \leq k \leq n \leq |V|$, $v_l \in V$ for all $0 \leq l \leq n$ and these vertices are all different. For any $i \in \Pi$, let $\mathcal{A}_i = (M, m_0, V, \delta, \nu)$ be the strategy automaton of player i , where:

- $M = \{v_0 v_0, v_0 v_1, \dots, v_{n-1} v_n, v_n v_k\} \cup \Pi \setminus \{i\}$.
As we want to be sure that the play ρ is followed by all players, we need to memorise which movement (edge) has to be chosen at each step of ρ . This is the role of $\{v_0 v_0, v_0 v_1, \dots, v_{n-1} v_n, v_n v_k\}$. But in case a player deviates from ρ , we only have to remember this player during the rest of the play (no matter if another player later deviates from ρ). This is the role of $\Pi \setminus \{i\}$.
- $m_0 = v_0 v_0$ (this memory state means that the play has not begun yet).
- $\delta : M \times V \rightarrow M$ is defined in this way: given $m \in M$ and $v \in V$,

$$\delta(m, v) := \begin{cases} j & \text{if } m = j \in \Pi \text{ or} \\ & (m = u_1 u_2, \text{ with } u_1, u_2 \in V, v \neq u_2 \text{ and } u_1 \in V_j), \\ v_l v_{l+1} & \text{if } m = u v_l \text{ for a certain } l \in \{0, \dots, n-1\}, u \in V, \\ & \text{and } v = v_l, \\ v_n v_k & \text{otherwise } (m = u v_n \text{ and } v = v_n). \end{cases}$$

Intuitively, m represents either a player to punish, or the edge that should, if following ρ , have been chosen at the last step of the current stage of the play, and v is the real last vertex of the current stage of the play.

Notice that in this definition of δ , j is different from i because if player i follows the strategy computed by this strategy automaton, one can be convinced that he does not deviate from the play ρ .

- $\nu : M \times V_i \rightarrow V$ is defined in this way: given $m \in M$ and $v \in V_i$,

$$\nu(m, v) := \begin{cases} \sigma_i^*(v) & \text{if } m = u_1 u_2 \text{ with } u_1, u_2 \in V \text{ and } v = u_2, \\ \sigma_{i,j}^*(v) & \text{if } m = j \in \Pi \text{ or} \\ & (m = u_1 u_2, \text{ with } u_1, u_2 \in V, v \neq u_2 \text{ and } u_1 \in V_j). \end{cases}$$

The idea is to play according to σ_i^* if everybody follows the play ρ , and switch to $\sigma_{i,j}^*$ if player j is the first player who has deviated from ρ .

Obviously, the strategy $\sigma_{\mathcal{A}_i}$ computed by the strategy automaton \mathcal{A}_i exactly corresponds to the strategy τ_i of the Nash equilibrium. And so, we can conclude that each strategy τ_i requires a memory of size at most $|M| \leq |\Pi| + |V|$. \square

D Example 3 continued

Example 21. Thanks to the proof of Theorem 13, we can construct a finite strategy automaton \mathcal{A}_1 that computes the strategy τ_1 of player 1 given in Example 12. The set M of memory states is $M = \{AA, AB, BC, CB\} \cup \{2\}$ since $\rho = A(BC)^\omega$, and the initial state is $m_0 = AA$. The memory update function $\delta : M \times V \rightarrow M$ and the transition choice function $\nu : M \times V_1 \rightarrow V$ are depicted in Figure 6: a label v/v' on an edge (m_1, m_2) means that $\delta(m_1, v) = m_2$, and $\nu(m_1, v) = v'$ if $v \in V_1$. If $v \notin V_1$, we indicate that ν does not return any advice by a ‘-’, and label the edge with $v/-$.

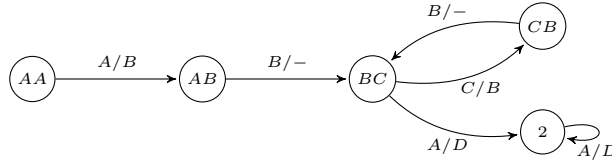


Fig. 6. The finite strategy automaton \mathcal{A}_1 .

E Sketch of proof of Theorem 14

Theorem 14 states that in every initialised multiplayer cost game that is cost-prefix-linear and *finite-memory* coalition-determined, there exists a Nash equilibrium with finite memory.

Proof (Sketch). The proof follows the same philosophy than the proof of Theorem 13 and keeps the same notations. Again we consider the Nash equilibrium $(\tau_i)_{i \in \Pi}$ defined in the proof of Theorem 10, whose outcome is $\rho := \langle (\sigma_i^*)_{i \in \Pi} \rangle_{v_0}$. We recall that for all $i \in \Pi$, the strategy τ_i depends on the strategies σ_i^* and $\sigma_{i,j}^*$ for $j \in \Pi \setminus \{i\}$. As the game \mathcal{G} is finite-memory coalition-determined by hypothesis, these strategies are assumed to be finite-memory. Given $i \in \Pi$ and $j \in \Pi \setminus \{i\}$, we denote by $\mathcal{A}^{\sigma_i^*}$ (resp. $\mathcal{A}^{\sigma_{i,j}^*}$) a finite strategy automaton for the strategy σ_i^* (resp. $\sigma_{i,j}^*$).

As in the proof of Theorem 13, each player needs to remember both the play ρ and who has to be punished. But here the play ρ is not anymore the outcome of a positional strategy profile: each σ_i^* is a finite-memory strategy. Nevertheless, in some sense, we can see the σ_i^* 's as positional strategies played on the product graph $G \times \mathcal{A}^{\sigma_1^*} \times \dots \times$

$\mathcal{A}^{\sigma_{\Pi}^*}$. This allows us to write $\rho := v_0 \dots v_{k-1}(v_k \dots v_n)^\omega$ where¹⁴ $0 \leq k \leq n \leq |V| \cdot \prod_{j \in \Pi} |\mathcal{A}^{\sigma_j^*}|$, $v_l \in V$ for all $0 \leq l \leq n$. Like in the proof of Theorem 13, we can now define, for any $i \in \Pi$, \mathcal{A}^{τ_i} , a finite strategy automaton for τ_i . In order to build explicitly \mathcal{A}^{τ_i} , we need to take into account, on one hand, the path ρ , and on the other hand, the memory of the punishing strategies $\sigma_{i,j}^*$. This enables to bound the size of \mathcal{A}^{τ_i} by $|V| \cdot \prod_{j \in \Pi} |\mathcal{A}^{\sigma_j^*}| + \sum_{j \in \Pi \setminus \{i\}} |\mathcal{A}^{\sigma_{i,j}^*}|$. \square

F Remark on the particular Min-Max cost games of Definition 15

Remark 22. Note that reachability-price and discounted-price games are zero-sum¹⁵ games, whereas the two other ones are not. For example, let us consider the average-price game \mathcal{G} depicted on Figure 7. The vertices of this game are A and B , and the number 0 or 1 associated to an edge corresponds with the price of this edge ($\pi(A, B) = \pi(B, B) = 1$ and the price of the other edges is zero).

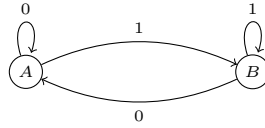


Fig. 7. Average-price game \mathcal{G} .

Let ρ be the play $ABAB^2A^2B^4A^4 \dots B^{2^n}A^{2^n} \dots$, where A^i means the concatenation of i A . Then the sequence of prices appearing along ρ is $101^20^21^40^4 \dots 1^{2^n}0^{2^n} \dots$, and so we get: $\text{AP}_{\text{Min}}(\rho) = \frac{2}{3}$ and $\text{AP}_{\text{Max}}(\rho) = \frac{1}{2}$. As these costs are not equal, the average-price game \mathcal{G} depicted on Figure 7 is not a zero-sum game. Since an average-price game is a special case of price-per-reward-average game, we can conclude that these two kinds of games are non zero-sum games.

G Part of the proof of Theorem 17

Proposition 23. *Let $\mathcal{G} = (\Pi, V, (V_i)_{i \in \Pi}, E, (\text{Cost}_i)_{i \in \Pi})$ be a multiplayer cost game where the cost function Cost_i belongs to $\{\text{RP}_{\text{Min}}, \text{DP}_{\text{Min}}, \text{AP}_{\text{Min}}, \text{PRAvg}_{\text{Min}}\}$ for each $i \in \Pi$. Then the game \mathcal{G} is cost-prefix-linear and positionally coalition-determined.*

Proof. Let \mathcal{G} be a multiplayer cost game where each cost function is RP_{Min} , DP_{Min} , AP_{Min} or $\text{PRAvg}_{\text{Min}}$. Let us first prove that the game \mathcal{G} is cost-prefix-linear. Given $j \in \Pi$, $v \in V$ and $hv \in \text{Hist}$, we consider the four possible cases for Cost_j . Let $\pi : E \rightarrow \mathbb{R}$ be a price function and $\vartheta : E \rightarrow \mathbb{R}$ be a diverging reward function. For the sake of simplicity, we write $hv := h_0 \dots h_k$ with $k \in \mathbb{N}$, $h_k = v$ and $h_l \in V$ for $l = 0, \dots, k$.

¹⁴ $|A|$ denotes the number of states of the automaton A .

¹⁵ Let us recall that a Min-Max cost game is zero-sum if and only if $\text{Cost}_{\text{Min}} = \text{Gain}_{\text{Max}}$.

Moreover, to avoid heavy notation, we do not explicitly show the dependency between Goal and j in the first case or between λ and j in the second case.

(i) Case $\text{Cost}_j = \text{RP}_{\text{Min}}$ for a given goal set $\text{Goal} \subseteq V$:

Let us distinguish two situations. If there exists $l \in \{0, \dots, k\}$ such that $h_l \in \text{Goal}$, then we set $a := \sum_{i=1}^n \pi(h_{i-1}, h_i) \in \mathbb{R}$ and $b := 0 \in \mathbb{R}^+$, where n is the least index such that $h_n \in \text{Goal}$. Let ρ be a play with $\text{First}(\rho) = v$, then it implies that $\text{RP}_{\text{Min}}(h\rho) = \sum_{i=1}^n \pi(h_{i-1}, h_i) = a + b \cdot \text{RP}_{\text{Min}}(\rho)$ (with the convention that $0 \cdot +\infty = 0$).

If there does not exist $l \in \{0, \dots, k\}$ such that $h_l \in \text{Goal}$, then we set $a := \sum_{i=1}^k \pi(h_{i-1}, h_i) \in \mathbb{R}$ and $b := 1 \in \mathbb{R}^+$. Let $\rho = \rho_0 \rho_1 \dots$ be a play such that $\text{First}(\rho) = v$. If $\text{RP}_{\text{Min}}(\rho)$ is infinite, then $\text{RP}_{\text{Min}}(h\rho) = +\infty = a + b \cdot \text{RP}_{\text{Min}}(\rho)$. Otherwise, if n is the least index in \mathbb{N} such that $\rho_n \in \text{Goal}$, then we have that:

$$\begin{aligned} \text{RP}_{\text{Min}}(h\rho) &= \sum_{i=1}^k \pi(h_{i-1}, h_i) + \sum_{i=1}^n \pi(\rho_{i-1}, \rho_i) \\ &= a + b \cdot \text{RP}_{\text{Min}}(\rho). \end{aligned}$$

(ii) Case $\text{Cost}_j = \text{DP}_{\text{Min}}(\lambda)$ for a given discount factor $\lambda \in]0, 1[$:

We set $a := (1 - \lambda) \sum_{i=1}^k \lambda^{i-1} \pi(h_{i-1}, h_i) \in \mathbb{R}$ and $b := \lambda^k \in \mathbb{R}^+$. Given a play $\rho = \rho_0 \rho_1 \dots$ such that $\text{First}(\rho) = v$ and $\eta := h\rho \in \text{Plays}$ (with $\eta = \eta_0 \eta_1 \dots$), we have that:

$$\begin{aligned} \text{DP}_{\text{Min}}(\lambda)(h\rho) &= \text{DP}_{\text{Min}}(\lambda)(\eta) \\ &= (1 - \lambda) \sum_{i=1}^{+\infty} \lambda^{i-1} \pi(\eta_{i-1}, \eta_i) \\ &= (1 - \lambda) \sum_{i=1}^k \lambda^{i-1} \pi(\eta_{i-1}, \eta_i) + (1 - \lambda) \sum_{i=k+1}^{+\infty} \lambda^{i-1} \pi(\eta_{i-1}, \eta_i) \\ &= (1 - \lambda) \sum_{i=1}^k \lambda^{i-1} \pi(h_{i-1}, h_i) + \lambda^k (1 - \lambda) \sum_{i=1}^{+\infty} \lambda^{i-1} \pi(\rho_{i-1}, \rho_i) \\ &= a + b \cdot \text{DP}_{\text{Min}}(\lambda)(\rho). \end{aligned}$$

(iii) Case $\text{Cost}_j = \text{AP}_{\text{Min}}$:

We set $a := 0 \in \mathbb{R}$ and $b := 1 \in \mathbb{R}^+$. Given $\rho \in \text{Plays}$ such that $\text{First}(\rho) = v$ and $\eta := h\rho \in \text{Plays}$ (with $\eta = \eta_0 \eta_1 \dots$), we show that:

$$\text{AP}_{\text{Min}}(h\rho) = \text{AP}_{\text{Min}}(\eta) = \text{AP}_{\text{Min}}(\rho).$$

If $\text{AP}_{\text{Min}}(\eta) = \text{AP}_{\text{Min}}(\rho) = +\infty$ or $-\infty$, the desired result obviously holds. Otherwise, let us set $x_n := \frac{1}{n} \sum_{i=1}^n \pi(\eta_{i-1}, \eta_i)$ and $y_n := \frac{1}{n} \sum_{i=1}^n \pi(\rho_{i-1}, \rho_i)$, for all $n \in \mathbb{N}_0$. By properties of the limit superior and definition of the AP_{Min} function, it holds that:

$$\limsup_{n \rightarrow +\infty} (x_n - y_n) \geq \text{AP}_{\text{Min}}(\eta) - \text{AP}_{\text{Min}}(\rho) \geq \liminf_{n \rightarrow +\infty} (x_n - y_n).$$

It remains to prove that the sequence $(x_n - y_n)_{n \in \mathbb{N}}$ converges to 0. For all $n > k$, we have that:

$$\begin{aligned} |x_n - y_n| &= \left| \frac{1}{n} \cdot \left(\sum_{i=1}^n \pi(\eta_{i-1}, \eta_i) - \sum_{i=k+1}^{k+n} \pi(\eta_{i-1}, \eta_i) \right) \right| \\ &= \frac{1}{n} \cdot \left| \sum_{i=1}^k \pi(\eta_{i-1}, \eta_i) - \sum_{i=n+1}^{n+k} \pi(\eta_{i-1}, \eta_i) \right|. \end{aligned}$$

As the absolute value is bounded independently of n (let us remind that E is finite), we can conclude that $(x_n - y_n)_{n \in \mathbb{N}}$ converges to 0, and so $\text{AP}_{\text{Min}}(\eta) = \text{AP}_{\text{Min}}(\rho)$.

(iv) Case $\text{Cost}_j = \text{PRAvg}_{\text{Min}}$:

We set $a := 0 \in \mathbb{R}$ and $b := 1 \in \mathbb{R}^+$. Given $\rho \in \text{Plays}$ such that $\text{First}(\rho) = v$ and $\eta := h\rho \in \text{Plays}$ (with $\eta = \eta_0\eta_1 \dots$), we show that:

$$\text{PRAvg}_{\text{Min}}(h\rho) = \text{PRAvg}_{\text{Min}}(\eta) = \text{PRAvg}_{\text{Min}}(\rho).$$

Thanks to several properties of \limsup , we have that:

$$\begin{aligned} \text{PRAvg}_{\text{Min}}(\rho) &= \limsup_{n \rightarrow +\infty} \frac{\sum_{i=1}^n \pi(\rho_{i-1}, \rho_i)}{\sum_{i=1}^n \vartheta(\rho_{i-1}, \rho_i)} \\ &= \limsup_{n \rightarrow +\infty} \frac{\sum_{i=1}^n \pi(\eta_{k+i-1}, \eta_{k+i})}{\sum_{i=1}^n \vartheta(\eta_{k+i-1}, \eta_{k+i})} \\ &= \limsup_{n \rightarrow +\infty} \frac{\sum_{i=1}^{n+k} \pi(\eta_{i-1}, \eta_i) - \sum_{i=1}^k \pi(\eta_{i-1}, \eta_i)}{\sum_{i=1}^{n+k} \vartheta(\eta_{i-1}, \eta_i) - \sum_{i=1}^k \vartheta(\eta_{i-1}, \eta_i)} \\ &= \limsup_{n \rightarrow +\infty} \frac{\sum_{i=1}^{n+k} \pi(\eta_{i-1}, \eta_i)}{\sum_{i=1}^{n+k} \vartheta(\eta_{i-1}, \eta_i) - \sum_{i=1}^k \vartheta(\eta_{i-1}, \eta_i)} \quad (8) \\ &= \limsup_{n \rightarrow +\infty} \frac{\sum_{i=1}^{n+k} \pi(\eta_{i-1}, \eta_i)}{\sum_{i=1}^{n+k} \vartheta(\eta_{i-1}, \eta_i)} \cdot \frac{1}{1 - \frac{\sum_{i=1}^k \vartheta(\eta_{i-1}, \eta_i)}{\sum_{i=1}^{n+k} \vartheta(\eta_{i-1}, \eta_i)}} \end{aligned}$$

$$\begin{aligned} &= \limsup_{n \rightarrow +\infty} \frac{\sum_{i=1}^{n+k} \pi(\eta_{i-1}, \eta_i)}{\sum_{i=1}^{n+k} \vartheta(\eta_{i-1}, \eta_i)} \quad (9) \\ &= \limsup_{n \rightarrow +\infty} \frac{\sum_{i=1}^n \pi(\eta_{i-1}, \eta_i)}{\sum_{i=1}^n \vartheta(\eta_{i-1}, \eta_i)} \\ &= \text{PRAvg}_{\text{Min}}(\eta) = \text{PRAvg}_{\text{Min}}(h\rho). \end{aligned}$$

Line (8) comes from the fact that the reward function ϑ is diverging, and from the following property: if $\lim_{n \rightarrow +\infty} b_n = b \in \mathbb{R}$, then $\limsup_{n \rightarrow +\infty} (a_n + b_n) = (\limsup_{n \rightarrow +\infty} a_n) + b$. Line (9) is implied by this property: if $\lim_{n \rightarrow +\infty} b_n = b > 0$, then $\limsup_{n \rightarrow +\infty} (a_n \cdot b_n) = (\limsup_{n \rightarrow +\infty} a_n) \cdot b$.

Note that, if the history h is empty, then $k = 0$ and, in all cases, a is equal to 0 and b to 1. This actually implies that $\text{Cost}_i(h\rho) = \text{Cost}_i(\rho)$ holds.

Let us now prove that the game \mathcal{G} is positionally coalition-determined. Given a player $i \in \Pi$, if $\text{Cost}_i = \text{RP}_{\text{Min}}$, then we take $\text{Gain}_{\text{Max}}^i = \text{RP}_{\text{Max}}$. We do the same for the other cases by defining the gain function $\text{Gain}_{\text{Max}}^i$ for the coalition as the counterpart

of Cost_i in Definition 15. Clearly, it holds that $\text{Cost}_i \geq \text{Gain}_{\text{Max}}^i$. Moreover, the Min-Max cost game $\mathcal{G}^i = (V, V_i, V \setminus V_i, E, \text{Cost}_i, \text{Gain}_{\text{Max}})$ is determined and has positional optimal strategies by Theorem 16. \square